BAB II

TINJAUAN PUSTAKA

2.1 Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Curah hujan satu milimeter artinya adalah dalam luasan satu meter persegi tempat yang datar, tertampung air setinggi satu milimeter atau tertampung air sebanyak satu liter. Intensitas hujan merupakan banyaknya curah hujan per-satuan jangka waktu tertentu. Jadi, apabila intensitas hujan dikatakan besar, itu tandanya hujan lebat dan dapat menimbulkan banjir. Berdasarkan infornasi yang diperoleh dari Badan Meteorologi dan Geofisika (BMG), tinggi curah hujan 1 mm sama dengan jumlah air hujan sebanyak 1 liter dalam luasan 1 meter persegi (1 mm = 1 liter/m2). Keadaan curah hujan dikatakan musim kering jika curah hujan kurang dari 50 mm/10 hari (< 50 mm/10 hari) dan musim hujan jika curah hujan mencapai lebih dari atau sarna dengan 50 mm/10 hari (≥ 50 mm/10 hari).Berdasarkan intensitas curah hujan dibedakan menjadi 3 yaitu hujan sedang berada diantara 20 dan 50 mm perhari, hujan lebat berada diantara 50 dan 100 mm perhari, dan hujan sangat lebat berada diatas 100 mm perhari.

2.2 Analisis Deret Waktu

Deret waktu adalah sekumpulan data yang berupa pengamatan yang diukur selama kurun waktu tertentu, berdasarkan waktu dengan interval yang sama. Dengan adanya deret waktu maka menjadikan adanya analisis deret waktu yang merupakan metode untuk melalukan peramalan. Peramalan deret waktu adalah penggunaan model untuk memprediksi nilai pada kejadian dimasa yang akan datang. Oleh karena itu data yang digunakan harus memiliki hubungan antara kejadian mendatang dengan kejadian di waktu sebelumnya (Ischak, 2018).

Dalam pemilihan metode peramalan terdapat faktor penting yaitu mengidentifikasi dan memahami pola data. Pengidentifikasian pola data dapat dilakukan dengan melihat plot data deret waktu yang disajikan sesuai dengan data. Pola data pada deret waktu terbagi menjadi 4 jenis, diantaranya (Makridakis S, *et al*, 1997) :

1. Pola Horizontal (H) yaitu pola data yang terjadi jika data berfluktuasi di sekitar nilai rata-rata konstan.
2. Pola Musiman (S) yaitu pola dataterjadi jika data dipengaruhi faktor musiman.
3. Pola Siklis (C) yaitu pola data yang terjadi jika data dipengaruhi fluktuasi ekonomi jangka panjang.
4. Pola Data Trend (T) yaitu pola data yang terjadi jika terjadi kenaikan ataupun penurunan jangka panjang pada data.

2.3 Analisis Kelompok

Salah satu teknik yang dapat digunakan untuke menyelesaikan masalah analisis data multivariat adalah teknik analisis *cluster*. Analisis *cluster* adalah salahsatu teknik pengelompokkan data ke dalam suatu kelompok berdasarkan karakteristik tertentu yang saling bebas sehingga individu-individu di dalam satu kelompok tersebut memiliki tingkat kemiripan yang tinggi terhadap satu sama lain, dan tidak mirip dengan individu lain pada kelompok yang berbeda (Nurcahyani, 2016 dalam Johnson, 1982).

2.4 Uji Korelasi Antar Lokasi

Uji korelasi digunakan untuk mengetahui keterkaitan antar lokasi dalam penelitian. Korelasi antar lokasi ke-i dan ke-j dapat dihitung dengan rumus :

$$\rho = \frac{\sum_{t=1}^{T}\left(z_{i,t} - \bar{z}_i\right)\left(z_{j,t} - \bar{z}_j\right)}{\left(\sum_{t=1}^{T}\left(z_{i,t} - \bar{z}_i\right)^2 \sum_{t=1}^{T}\left(z_{j,t} - \bar{z}_j\right)^2\right)}$$

Keterangan :

ρ : kofesiensi korelasi

$z_{i,t}$: nilai pengamatan pada waktu ke-t untuk lokasi ke-i

$z_{j,t}$: nilai pengamatan pada waktu ke-t untuk lokasi ke-j

Dua variabel dikatakan berkorelasi apabila perubahan pada variabel yang satu akan diikuti perubahan pada variabel yang lain secara teratur dengan arah yang sama (positif) atau berlawanan (negatif). Koefisien korelasi berkisar antara 1 sampai dengan -1 menurut Makridakis dkk., 1999. Apabila $\rho = 0$ artinya dua lokasi tidak saling berkorelasi, $\rho < 0$ artinya dia lokasi saling berkorelasi negatif dan $\rho > 0$ artinya dua lokasi saling berkorelasi positif. Nilai koefisien korelasi mendekati 1 menunjukan adanya korelasi yang sempurna.

2.5 Uji Heterogenitas Lokasi

Menurut Karlina, et al., (2014) metode indeks Gini merupakan rasioanalisis yang sangat merepresentatif data dalam masyarakat yang heterogen. Metode indeks Gini biasanya digunakan untuk mengetahui tingkat pemerataan pendapatan masyarakat dengan melihat nilai indeks Gini yang dibagi menjadi beberapa kriteria diantaranya Gn = 0 artinya pemerataan sempurna dan Gn = 1 artinya pemerataan tidak sempurna. Indeks Gini merupakan suatu metode yang dapat digunakan untuk membandingkan dari satu waktu ke waktu atau dari satu lokasi ke lokasi yang lain. Berikut uji hipotesis heterogenitas.

$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ (terdapat homogenitas spasial)

$H_1 : \sigma_1^2 \neq \sigma_1^2$ (paling tidak terdapat satu heterogenitas spasial)

Kriteria Uji :

H_0 ditolak jika nilai G lebih dari sama dengan 1.

Statistik uji :

$$G = 1 + \frac{1}{n} - \frac{2}{n^2 \bar{y}_i} \times \sum_{i=1}^{n_i} y_i$$

dengan

G : indeks Gini,

n : jumlah data,

n_i : jumlah data pada lokasi ke-i,

$\bar{y}_i$: merupakan rata-rata dari masing-masing variabel yang diamati.

2.6 Stasioneritas

Syarat penting dalam pemodelan data deret waktu adalah kestasioneran data. Kestasioneran data dapat dilihat dari distribusi peluang bersama dari observasi $y_t, y_{t+1}, \dots, y_{t+n},$ sama persis dengan distribusi peluang bersama dari observasi $y_{t+k}, y_{t+k+1}, \dots, y_{t+k+n},$ (Montgomery et al. 2015). Suatu data deret waktu dikatakan stasioner jika nilai rataan dan ragam konstan serta tidak terdapat pola musiman. Stasioner dalam rataan dan ragam dapat diperiksa secara eksplorasi yaitu dilihat dari plot antara nilai observasi dengan waktu apakah polanya sudah konstan atau belum. Selain itu, kestasioneran dalam rataan dapat dilihat juga dengan uji formal yaitu uji *augmented Dickey Fuller* (ADF).

Uji ADF merupakan pengujian kestasioneran secara formal dengan melihat apakah data deret waktu mengandung akar unit (*unit root*). Uji *Augmented Dickey Fuller* (ADF) merupakan uji akar-akar unit metode *Dickey Fuller* untuk model *autoregressive* tingkat p dengan $p \geq 1$. Persamaan model nilai uji uni root dijelaskan dengan model nilai rataan nol (*zero mean*) dengan persamaan sebagai berikut (Rusdi 2011) :

$$\nabla \mathrm{y}(t) = a_0 + a_1 t + \gamma y(t-1) + \sum_{i=2}^{p} \beta_i y(t-i+1) + e(t)$$

dengan,

$\nabla \mathrm{y}(t)$: selisih data pengamatan waktu ke-*t* dengan waktu sebelumnya

$\mathrm{y}(t)$: nilai deret waktu ke-*t*

a_0 : nilai itersep

a_1 : koefesien regresi untuk *trend*

γ : koefesien regresi untuk lag y

β_i : koefesien regresi untuk pembedaan lag y

$e(t)$: nilai sisaan pada waktu ke-*t*

Hipotesis berdasarkan akar unit (Wei 2015) :

$H_0 : \gamma = 1$ (data mengandung akar unit)

$H_1 : \gamma \neq 1$ (data tidak mengandung akar unit)

Dengan nilai statistik uji,

$$ADF = t_{hitung} = \frac{\hat{\gamma} - 1}{S_{\hat{\gamma}}}$$

Hipotesis nol ditolak jika nilai statistik ADF $\left(t_{hitung}\right)$ lebih besar dari nilai kritis Tabel MacKnnon sehingga dapat disimpulkan bahwa data tidak mengandung akar unit atau data sudah stasioner. Jika tidak stasioner dalam rataan, dilakukan pembedaan dan jika data tidak stasioner dalam ragam, dilakukan transformasi untuk menstabilkan ragam.

2.7 Identifikasi Model Ruang Waktu (*Space Time*)

Identifikasi model ruang waktu meliputi identifikasi terhadap ordo waktu dan ordo spasial. Ordo waktu dapat diidentifikasi dengan *matrix autocorrelation function* (MACF) dan *matrix partial autocorrelation function* (MPACF). Ordo spasial dibatasi pada ordo satu karena semakin tinggi ordo semakin sulit dalam menginterpretasikan (Setiawan 2015).

2.5.1 *Matrix Autocorrelation Function (MACF)*

Jika diberikan vektor deret waktu sebanyak n observasi $Z_1, Z_2, \ldots, Z_n$, dapat dihitung fungsi matriks korelasi sampel sebagai berikut:

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{n-k}\left(Z_{i,t} - \bar{Z}_i\right)\left(Z_{j,t-k} - \bar{Z}_j\right)}{\left[\sum_{t=1}^{n}\left(Z_{i,t} - \bar{Z}_i\right)^2 \sum_{t=1}^{T}\left(Z_{j,t-k} - \bar{Z}_j\right)^2\right]^{\frac{1}{2}}}$$

Dengan,

i : lokasi-*i*

j : lokasi-*j*

$\hat{\rho}_{ij}(k)$: sampel korelasi silang komponen deret waktu lokasi-*i* dan lokasi-*j*

$\bar{Z}_i$: rata-rata sampel pada vektor deret waktu lokasi-*i*

$\bar{Z}_j$: rata-rata sampel pada vektor deret waktu lokasi-*j*

n : banyaknya data pengamatan

k : lag waktu

Matrix autocorrelation function (MACF) sangat berguna untuk mengidentifikasi ordo model MA jika matriks korelasinya nol diluar proses lag q pada vektor MA (q). Matriks menjadi rumit jika dimensinya bertambah dan polanya sulit untuk dikenali. Masalah tersebut dapat diatasi dengan metode yang sederhana untuk meringkas korelasi sampel yaitu dengan simbol $(+), (-), (.)$ pada matriks korelasi sampel ke (i, j). Simbol $(+)$ diartikan sebagai $\hat{\rho}_{ij}(k)$ lebih besar dari 2 kali standar error dan menunjukkan korelasi positif, simbol $(-)$ diartikan sebagai $\hat{\rho}_{ij}(k)$ kurang dari -2 kali standar error dan menunjukan korelasi negatif, dan simbol $(.)$ diartikan sebagai $\hat{\rho}_{ij}(k)$ berada di $\pm$ 2 kali standar error dan menunjukkan tidak adanya korelasi (Wei 2015).

2.5.2 *Matrix Partial Autocorrelation Function (MPACF)*

Matrix partial autocorrelation function (MPACF) sangat diperlukan dalam mengidentifikasi model AR. Korelasi antara $Z(t)$ dan $Z(t+k)$ dapat diketahui setelah ketergantungan linier variabel $Z(t+1), \ldots, Z(t+k-1)$ dihilangkan. Sebagaimana dirumuskan sebagai berikut (Wei 2015) :

$$\emptyset_{kk} = \frac{Cov\left[\left(Z_t - \hat{Z}_t\right), \left(Z_{t+k} - \hat{Z}_{t+k}\right)\right]}{\sqrt{Var\left(Z_t - \hat{Z}_t\right)}\sqrt{Var\left(Z_{t+k} - \hat{Z}_{t+k}\right)}}$$

dengan,

$\hat{Z}(t)$ dan $\hat{Z}(t+k)$: nilai MSE terkecil penduga regresi linier

$\emptyset_{kk}$: matriks fungsi korelasi parsial pada lag waktu ke-k

2.8 Model *Generalized Space Time Autoregressive* (GSTAR)

Model *generalized space time autoregressive* (GSTAR) pertama kali diperkenalkan oleh Borovkova, Lopuha dan Ruchjana pada tahun 2002 dalam Wutsqa dkk, (2010). Model *generalized space time autoregressive* (GSTAR) merupakan perluasan model *space time autoregressive* (STAR). Model *generalized space time autoregressive* (GSTAR) merupakan pengembangan dari model *space time autoregressive* (STAR) dengan asumsi parameter-parameter model berubah untuk setiap lokasi sehingga model *generalized space time autoregressive* (GSTAR) cenderung lebih fleksibel dibandingkan model *space time autoregressive* (STAR). Model GSTAR$(p, \lambda_1, \ldots, \lambda_i)$ apat dituliskan sebagai berkut :

$$\boldsymbol{z}_{(t)} = \sum_{k=1}^{p} \left[\emptyset_{\boldsymbol{k0}} + \sum_{l=1}^{\lambda_s} \emptyset_{\boldsymbol{kl}} \boldsymbol{W}^{(\boldsymbol{l})} \right] \boldsymbol{z}_{(t-k)} + \boldsymbol{e}_{(t)}$$

dimana :

$\boldsymbol{z}_{(t)}$ = vektor acak berukuran $N \times l$ pada waktu t

λ_s = orde spasial ke-s dari bentuk autoregresif

p = orde autoregresif

$\boldsymbol{\phi}_{\boldsymbol{kl}}$ = parameter autoregresif pada lag waktu k dan lag spasial l

$\boldsymbol{W}^{(l)}$ = matriks pembobot berukuran $N \times N$ untuk lag spasial $l = 0,1,\dots,\lambda_s$

$\boldsymbol{e}_{(t)}$ = vektor galat berukuran $N \times l$ bersifat *white noise* dan berdistribusi normal multivariat

2.9 Model *Generalized Space Time Autoregressive Integrated* (GSTAR-I)

Model *generalized space time autoregressive integrated* (GSTAR-I) merupakan model dengan parameter yang bervariasi menurut lokasi. Model *generalized space time autoregressive integrated* (GSTAR-I) merupakan model perluasan dari model *generalized space time autoregressive* (GSTAR) yang digunakan pada data tidak stasioner. Model *generalized space time autoregressive integrated* (GSTAR-I) deiperkenalkan oleh Min dkk (2010) dengan model ditulisakan sebagai berikut :

$$z^*_{(t)} = \sum_{k=1}^{p} \left[\emptyset_{k0} W^{(0)} + \sum_{l=1}^{\lambda_s} \emptyset_{kl} W^{(l)} \right] z^*_{(t-k)} + e_{(t)}$$

dimana :

$z^*_{(t)}$: vektor dari lokasi N pada waktu ke-t

λ_s : orde spasial ke-*s*

p : orde autoregresif

$\boldsymbol{W}^{(l)}$: matriks pembobot berukuran $N \times N$ untuk spasial lag 1

ϕ_{kl} : parameter autoregresif pada lag waktu k danlag spasila l

$e_{(t)}$: vektor galat berukuran $N \times 1$ bersifat *white noise* dan berdistribusi normal multivariat

2.10 Bobot Lokasi

Bobot lokasi berfungsi untuk menggambarkan hubungan spasial antar wilayah. Bobot lokasi yang sesuai dengan penelitian ini adalah bobot lokasi invers jarak, penentuan ini mempertimbangkan lokasi penelitian yang digunakan adalah berupa titik dan lokasi yang berdekatan diasumsikan memiliki keterkaitan yang erat.

Bobot invers jarak didapatkan dari penghitungan berdasarkan jarak sebenarnya antar lokasi. Lokasi yang berdekatan mendapatkan nilai bobot yang lebih besar dan lokasi yang berjauhan mendapatkan nilai bobot yang lebih kecil (Faizah dan Setiawan 2013, dalam Dwi 2018). Pembobot ini dihitung dari garis lintang dan garis bujur koordinat jarak yang diamati. Persamaan d_{ij} sebagai berikut :

$$d_{ij} = d_{ij} = \left(\left[x_i(u_i) - x_j(u_j)\right]^2 + \left[x_i(v_i) - x_j(v_j)\right]^2\right)^{1/2}$$

$$w_{ij} *= w_{ij} *= {}^{1}/_{d_{ij}} = {}^{1}/_{d_{ij}}$$

$$w_{ji} = \frac{w_{ij} *}{\sum_j w_{ij} *}$$

Dengan,

i : lokasi-*i*

j : lokasi-*j*

x_i : indeks lokasi-*i*

x_j : indeks lokasi-*j*

u_i : koordinat lintang lokasi-*i*

v_i : koordinat lintanglokasi-*j*

d_{ij} : jarak Euclid antara lokasi-*i* dan lokasi-*j*

w_{ji} : bobot kebalikan jarak lokasi-*i* dan lokasi-*j*

$w_{ij}*$: elemen dari baris ke-*i*, kolom ke-*j* pada matriks bobot kebalikan jarak

2.11 Uji Diagnostik Model

Diagnostik model bertujuan untuk membuktikan bahwa model yang didapatkan layak digunakan dalam peramalan. Model *time series multivariate* memiliki dua asumsi yang harus terpenuhi yang harus terpenuhi, yaitu bersifat *white noise* dan berdistribusi normal *multivariate*.

2.9.1 Uji *White Noise*

Uji multivariate white noise bertujuan untuk melihat apakah residual dari model sudah saling independen (bebas) antara satu dengan yang lain. Pengujian asumsi ini dapat dilakukan dengan melihat plot MACF dari residual. Apabila pada plot MACF tidak terdapat lag yang signifikan bebarti residual bersifat white noise. Namum, selain melalui plot MACF dapat juga dilakukan pengujian hipotesis dengan uji *Portmanteau* yang merupakan generasi dari uji L-jung Box untuk kasus multivariate. Uji ini pertama kali diperkenalkan oleh BOX dan Pierce pada tahun 1970. Berikut adalah hipotesis pengujian :

H_0 : $\rho_1 = \rho_2 = \cdots = \rho_m = 0$ (residual bersifat *white nosie*)

H_1 : minimal terdapat satu $\rho_1 = \rho_2 = \cdots = \rho_m \neq 0$ (residual tidak bersifat *white noise*)

Statistik uji :

$$Q_N(m) = T^2 \sum_{t=1}^{m} \frac{1}{T-t} tr\left(\hat{\Gamma}_t' \hat{\Gamma}_0^{-1} \hat{\Gamma}_t' \hat{\Gamma}_0^{-1}\right)$$

dengan

$\hat{\Gamma}_t'$: matriks kovarian silang pada waktu t

N : banyak peubah (lokasi)

M : lag waktu ke-m

T : banyaknya pengamatan

Kriteria keputusan terima H_0 jika stastistik uji $Q_N(m) < \chi^2_{N^2m}$ atau $p - value >$ taraf signifikansi (α) yang berarti residual bersifat *white noise*.

2.12 Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan untuk memilih terbaik dari beberapa model yang dinyatakan layak pada uji diagnostik. Penentuan model berdasarkan skema plot MACF dan MPACF secara teori tidak prktis karena tergantung pada pengalama peneliti. Salah satu kriteria pemilihan dalam penentuan model terbaik dengan menggunakan AIC (*Akaike's Information Criterion*). Model dikatakan baik jika nilai AIC terkecil dari beberapa model deret waktu. Rumus untuk mendapatkan AIC (Wei 2006) :

$$AIC_p = ln|S(p)| + \frac{2pm^2}{n}$$

dimana :

m : banyaknya parameter yang diduga dalam model

n : banyaknya pengamatan

$S(p)$: jumlah kuadrat sisaan

p : orde model VAR

2.13 Ketepatan Model

Dalam melakukan suatu peramalan yang merupakan kegiatan memprediksi masa depan dengan menggunakan data di masa lampau, hasil yang akan didapatkan tidaklah sama dengan data yang sesugguhnya (Ishcak, 2018). Maka dari itu usaha untuk membuat nilai error seminimal munggkin dibutuhkan pada proses peramalan.

Salah satu tingkat akurasi peramalan dapat diukur daari nilai *Mean Absolute Persentage Error* (MAPE) yaitu rata-rata persentase kesalahan pertama dari beberapa periode. Tingkat keakuratan dapat dijelaskan dengan membandingkan nilai yang diproyeksikan dengan nilai *actual*. Untuk melakukan peramalan dan untuk mengetahui akuratnya sebuah model maka nilai akurasinya harus semakin kecil.

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{Y_t - \widehat{Y}_t}{Y_t} \right|}{n} \times 100\%$$

Untuk ketentuan MAPE adalah sebagai berikut :

Tabel 2. 1 Tabel nilai MAPE

MAPE	Keterangan
$< 10\%$	Sangat Baik
$< 20\%$	Baik
$< 30\%$	Cukup Baik
$> 30\%$	Tidak Akurat

Sumber : (Lewis, 1997)